The inefficacy of superficial similarities for establishing instructor-student relationships*

Reproducing 'Taking It to the Next Level'

Amy Farrow

April 9th, 2021

Abstract

This paper reproduces the 2019 paper 'Taking It to the Next Level', which evaluates an intervention to improve instructor-student relationship. The modeling results indicate that the intervention, consisting of informing instructors and students about commonalities, has a weak positive effect on student perceptions of instructor-student similarity, but no effect on student perceptions of instructor-student relationship, instructor perception of similarity or instructor-student relationship, grades, or re-enrollment. While the scalability and affordability of the intervention are desirable, there are no results in any of the targeted measures: those that affect and reflect college retention. These results are consistent with the original paper.

Keywords: instructor-student relationship, college, reproduction study

Contents

*Code and data are available at: github.com/amycfarrow/takingittothenextlevelrepro.

1 Introduction

College is perceived to be a meritocratic tool for social mobility and career success (Bowen, Chingos, and McPherson 2009). Unfortunately, retention is a large problem in contemporary American colleges, and many students begin degrees without completing them; six-year completion rates for full-time first-time students range from 51% to 86%, depending on school (Delbanco 2015). Even when controlling for pre-college test scores and initial enrollments, completion rate disparities exist based on parental education, socio-economic status, and race/ethnicity (Delbanco 2015). Disparities in college completion lead to further entrenchment of long-standing inequalities (Delbanco 2015). Thus, measures to help students persist in college completion are desirable to reduce wasted resources, and to increase societal equity.

Many interventions to increase completion rates have been suggested and tested. For example, Evans et al. (2020) studied case managment techniques; Gurantz, Hurwitz, and Smith (2017) investigated the effect of high school recruitment; Bers and Schuetz (2014) identify the importance of interpersonal relationships; Gilbert and Horn (2020) argue that certificates awarded prior to degrees can ameliorate the 'college-completion crisis'. These are only a small sample of the interventions suggested. Implementing an effective program is difficult due to complicated causes of attrition, embedded social inequalities, and expenses.

Robinson, Scott, and Gottfried (2019)'s paper, "Taking It to the Next Level: A Field Experiment to Improve Instructor-Student Relationships in College", tests one intervention. In this field experiment, they tested the effect of instructor-student similarity on instructor-student relationships and measures of student success. Based on extensive K-12 research about the importance of instructor-student relationship for student success, Robinson, Scott, and Gottfried (2019) aimed to establish how instructor-student relationships could be improved at the college level, and to test if this improvement had a positive result.

This paper reproduces Robinson, Scott, and Gottfried (2019)'s original study, using anonymized data provided by the authors. Linear, logistic, and ordinal logistic models are used to predict outcomes including perceived ISR, grade, and re-enrollment. These models show that treatment slightly improves student perception of instructor-student similarity, but does not significantly affect student or instructor ISR perception, grade, exam grade, or future enrollment. In addition to the models in the original paper, this paper explores models using different controls and models that consider specific subgroups of the study participants.

2 Methods

Robinson, Scott, and Gottfried (2019) conducted a randomized controlled trial to assess the impact of awareness of instructor-student similarities on perceived similarity, instructor-student relationship, course grade, and re-enrollment.

2.1 Participants

The study took place in the 2017 spring semester at a large Californian University. The study included 120 instructors and their 2,749 students. The instructors participated in the study based on interest and a gift-card incentive, and their students were invited to participate unincentivized.Students were only enrolled in the study for one class, in the event that they were taking classes with multiple participating instructors.

2.2 Treatment and control

Participating students were randomly assigned to either treatment or control. At the beginning of the term, all participating students and instructors were given "get to know you" surveys. Using those responses, for each student in the treatment group, seven commonalities were identified between student and instructor (for example, perhaps both student and instructor binge-watch TV to relieve stress, or appreciate loyalty as the most important friend quality), and both student and instructor were informed of these commonalities. They completed a few questions about the similarities and were reminded of them through the semester to ensure they were internalized. Students in the control group were informed about similarities they shared with students in another part of the country, and instructors were told nothing about these students.

2.3 Procedures

All students participated in a survey immediately following the treatment or the placebo. They were surveyed again at the end of the course. Instructors were surveyed only at the end of the course.

2.4 Measures

Robinson, Scott, and Gottfried (2019) identify key measures. Some are extracted from the student survey at the beginning of the term:

- s1_sim: Immediately after the treatment or placebo, students answered six questions about their perceived similarity to the instructor, on scales of 1 to 5. These responses were averaged to create a similarity scale.
  - Overall, how similar to your instructor's values do you think your values are?
  - How similar are your goals for the course and your instructor's goals?
  - In general, how similar do you think your views about the course content and your instructor's are?
  - How much do you think you have in common with your instructor?
  - How similar do you think your personality is compared to your instructor's?
  - Overall, how similar do you think you and your instructor are?
- s1_tsr: Immediately after the treatment or placebo, students answered seven questions about their perception of the instructor-student relationship, on scales of 1 to 5. These responses were averaged to create an ISR scale.
  - How much do you think you will enjoy learning from this instructor?
  - How friendly do you think this instructor will be towards you?
  - How encouraging do you think this instructor will be towards you?
  - If you came back to visit this instructor three years from now, how excited do you think they would be?
  - How motivating do you think you will find this instructor's class?
  - How caring do you think this instructor will be towards you?
  - Overall, how much do you think you will learn from this instructor?
- s1_female: The student's gender.

Others are extracted from the student survey at the end of the term:

- s2_sim: At the end of the semester, students answered the above six similarity scale questions again.
- s2_tsr: At the end of the semester, students answered the above seven ISR scale questions again.

Others come from the instructor survey at the end of the term:

- t2_sim1: At the end of the semester, instructors answered only one quesiton about similarity with the student, on a scale of 1 to 5.
  - Overall, how similar do you think you and STUDENTNAME are?
- t2_tsr: At the end of the semester, instructors answered seven questions about their perception of the instructor-student relationship, on scales of 1 to 5. These responses were averaged to create an ISR scale.
  - How much did you enjoy helping STUDENTNAME learn?
  - How caring was STUDENTNAME towards you?
  - How often did you say something encouraging to STUDENTNAME?
  - How friendly was STUDENTNAME towards you?
  - If this student came back to visit you three years from now, how excited would you be?
  - How motivating did STUDENTNAME find the activities that you plan for class?
  - Overall, how much did STUDENTNAME learn from you?
- t2_finalexam: Instructors were asked to report the student's grade on their final exam, paper, or project.

Finally, some are extracted from the university's internal records:

Table 1: Demographics for treatment and control groups

treatment	N	Missing		Control	Treatment	Total	p
				N(%) = 1388 (50.5)	N(%) = 1361 (49.5)	N(%) = 2749	
s_female	2740	9	0	552 (39.9)	536 (39.5)	1088 (39.7)	0.855
			1	831 (60.1)	821 (60.5)	1652 (60.3)	
t_female	2749	0	0	298 (21.5)	287 (21.1)	585 (21.3)	0.843
			1	1090 (78.5)	1074 (78.9)	2164 (78.7)	
s_race	2741	8	1	309 (22.3)	281 (20.7)	590 (21.5)	0.259
			2	71 (5.1)	57 (4.2)	128 (4.7)	
			3	687 (49.7)	725 (53.4)	1412 (51.5)	
			4	166 (12.0)	149 (11.0)	315 (11.5)	
			5	3 (0.2)	0 (0.0)	3 (0.1)	
			6	5 (0.4)	2 (0.1)	7 (0.3)	
			7	44 (3.2)	43 (3.2)	87 (3.2)	
			8	98 (7.1)	101 (7.4)	199 (7.3)	
t_race	2749	0	1	931 (67.1)	911 (66.9)	1842 (67.0)	0.998
			2	51 (3.7)	51 (3.7)	102 (3.7)	
			3	82 (5.9)	82 (6.0)	164 (6.0)	
			4	212 (15.3)	213 (15.7)	425 (15.5)	
			6	28 (2.0)	25 (1.8)	53 (1.9)	
			8	84 (6.1)	79 (5.8)	163 (5.9)	
s1_age	2633	116	Mean (SD)	22.4 (5.1)	22.4 (5.0)	22.4 (5.1)	0.927
t_age	2738	11	Mean (SD)	44.0 (10.7)	43.9 (10.7)	44.0 (10.7)	0.815
s_firstgen	2733	16	0	778 (56.3)	778 (57.5)	1556 (56.9)	0.549
			1	603 (43.7)	574 (42.5)	1177 (43.1)	
t_firstgen	2749	0	0	1066 (76.8)	1042 (76.6)	2108 (76.7)	0.918
			1	322 (23.2)	319 (23.4)	641 (23.3)	
ir_f16_gpa	2577	172	Mean (SD)	3.0 (0.7)	3.0 (0.7)	3.0 (0.7)	0.756
ir_t_s17enrltot	2641	108	Mean (SD)	111.2 (69.6)	110.9 (69.3)	111.1 (69.4)	0.912

- grade: The final grade that the student received in the course.
- std_grade: The student's final grade, standardized against other grades in the course.
- ir_f16_gpa: The student's cumulative GPA after the Fall 2016 term.
- f17_enrolled: The student's status as of Fall term 2017: not enrolled or enrolled.

TODO: evaluate scales' validity

3 Data

TODO: add all citations for R packages

Table 1 shows the demographics for the treatment and control samples. There are no significant differences between the treatment and control groups.

However, there are differences between the students and instructors. Notably, the student sample is 60.3% female, while the instructor sample is 78.7% female. The student sample is 21.5% White/Caucasian and 51.5% Hispanic American or Latino/a, while the instructor sample is 67.0% White/Caucasian and 6.0% Hispanic American or Latino/a. 43.1% of the students are first generation university attendees, while only 23.3% of the instructors are.

TODO: graphs showing instructor vs. student demographics. Implications.

Table 2 shows summary statistics for key variables identified by Robinson, Scott, and Gottfried (2019), including missing values.

Table 2: Descriptive statistics for treatment and control groups

treatment	N	Missing		Control	Treatment	Total
				N(%) = 1388 (50.5)	N(%) = 1361 (49.5)	N(%) = 2749
s1_sim	2658	91	Mean (SD)	3.5 (0.7)	3.6 (0.7)	3.5 (0.7)
s1_tsr	2749	0	Mean (SD)	4.1 (0.6)	4.1 (0.6)	4.1 (0.6)
s2_sim	2106	643	Mean (SD)	3.5 (0.8)	3.6 (0.8)	3.6 (0.8)
s2_tsr	2106	643	Mean (SD)	4.1 (0.8)	4.1 (0.7)	4.1 (0.8)
t2_sim1	2548	201	Mean (SD)	2.7 (1.1)	2.7 (1.0)	2.7 (1.1)
t2_tsr	2564	185	Mean (SD)	3.5 (0.9)	3.5 (0.9)	3.5 (0.9)
grade	2538	211	Mean (SD)	3.0 (1.1)	3.0 (1.1)	3.0 (1.1)
std_grade	2538	211	Mean (SD)	-0.0 (1.0)	0.0 (1.0)	0.0 (1.0)
t2_finalexam	2416	333	Mean (SD)	3.0 (1.1)	3.0 (1.1)	3.0 (1.1)
f17_enrolled	2190	559	0	91 (8.3)	97 (8.9)	188 (8.6)
			1	1010 (91.7)	992 (91.1)	2002 (91.4)
s1_female	2615	134	0	533 (40.3)	510 (39.4)	1043 (39.9)
			1	788 (59.7)	784 (60.6)	1572 (60.1)

4 Models

4.1 Reproduction models

$treatment_i$ is the indicator that treatment was given.

X_{1i} is a vector of student-level covariates (pre-intervention measures included).

X_{2j} is a vector of instructor-level covariates.

ϵ_{ij} is a clustered residual.

β_0, $beta_1$, Γ_1, Γ_2, and a_k are coefficients on the resulting models.

4.1.1 Linear models

Equation (1) is Robinson, Scott, and Gottfried (2019)'s linear model, used for continuous outcomes, which include complete scale outcomes (immediate student similarity rating (s1_sim), end of semester student similarity rating (s2_sim), end of semester student ISR rating (s2_tsr), and end of semester instructor ISR rating (t2_tsr)) and grade-based outcomes, measured in GPA (course grade (grade) and objectively graded exam grade (t2_finalexam)).

$$Outcome_{ij} = \beta_0 + \beta_1 treatment_i + X_{1i}\Gamma_1 + X_{2j}\Gamma_2 + \epsilon_{ij} \tag{1}$$

4.1.2 Ordinal logistic models

Equation (2) is Robinson, Scott, and Gottfried (2019)'s ordinal logistic model, used for the ordinal outcome, which is end of semester instructor similarity rating (t2_sim1).

$$prob(outcome_{ij}) = a_k + \beta_1 treatment_i + X_{1i}\Gamma_1 + X_{2j}\Gamma_2 + \epsilon_{ij} > k \tag{2}$$

4.1.3 Logistic models

Equation (3) is Robinson, Scott, and Gottfried (2019)'s logistic model, used for the binary outcome, which is enrollment in Fall term 2017 (f17_enrolled).

Table 3: Descriptive statistics and correlation matrix for key continuous variables

	N	Missing	Mean (SD)	s1_sim	s1_tsr	s2_sim	s2_tsr	t2_sim1	t2_tsr	grade	std_grade	t2_finalexam	ir_f16_gpa
s1_sim	2658	91	3.5 (0.7)	-									
s1_tsr	2749	0	4.1 (0.6)	0.67***	-								
s2_sim	2106	643	3.6 (0.8)	0.55***	0.49***	-							
s2_tsr	2106	643	4.1 (0.8)	0.41***	0.55***	0.79***	-						
t2_sim1	2548	201	2.7 (1.1)	0.14***	0.15***	0.23***	0.24***	-					
t2_tsr	2564	185	3.5 (0.9)	0.18***	0.2***	0.27***	0.31***	0.68***	-				
grade	2538	211	3.0 (1.1)	0.06**	0.11***	0.2***	0.23***	0.39***	0.46***	-			
std_grade	2538	211	0.0 (1.0)	0.01	0.02	0.11***	0.11***	0.31***	0.36***	0.81***	-		
t2_finalexam	2416	333	3.0 (1.1)	0.05*	0.1***	0.16***	0.18***	0.39***	0.45***	0.81***	0.63***	-	
ir_f16_gpa	2577	172	3.0 (0.7)	0.03	0.02	0.05*	0.05*	0.18***	0.2***	0.44***	0.42***	0.38***	-

$$prob(outcome_{ij}) = a_k + \beta_1 treatment_i + X_{1i}\Gamma_1 + X_{2j}\Gamma_2 + \epsilon_{ij} > k \tag{3}$$

4.2 Additional models

TODO: additional models Consider: More covariates: race, class size, first generation New covariates: instructor-student commonalities (gender, race) Models looking at subgroups of the sample: Only those with matching traits, only small classes Focus on grade as the outcome.

5 Results

5.1 Reproduction results

Table 3 displays a correlation matrix for the measures that Robinson, Scott, and Gottfried (2019) identified as significant. Some measures, unsurprisingly, are highly correlated: grade and standardized grade, and grade and final exam grade. Others are moderately correlated: student initial perceptions of similarity and instructor-student relationship, student end-of-term perceptions of similarity and instructor-student relationship, and instructor end-of-term perceptions of similarity and instructor-student relationship. These last three pairs could indicate that the similarity perception scales and instructor-student relationship perception scales may not be measuring distinct concepts, or that feelings of similarity and positivity in relationships are strongly associated in the classroom.

In Table 4, Model 1 shows a statistically significant relationship between treatment and student initial similarity perception. However, the expected increase is only 0.155 on a scale of 1 to 5. Model 2 shows a statistically significant relationship between treatment and student end-of-semester similarity perception. However, the expected increase is only 0.095 on a scale of 1 to 5.

Model 3 shows no significant relationship between treatment and student end-of-semester ISR perception. However, it does show a significant correlation between anticipated student ISR and student end-of-semester ISR. For every 1 point increase in anticipated ISR on a 1 to 5 scale, the expected increase in end-of-semester ISR is 0.674.

Models 4, 5, 6, 7, and 8 show no significant relationship between treatment and instructor end-of-semester ISR perception, course grade, objectively graded exam grade, instructor perception of similarity at the end of the term, and enrollment in the subsequent semester.

TODO: clustered residuals

5.2 Additional results

TODO: additional results

Table 4: Model results

	Dependent variable:							
	s1_sim	s2_sim	s2_tsr	t2_tsr	grade	t2_finalexam	t2_sim1	f17_enrolled
	OLS	*OLS*	*OLS*	*OLS*	*OLS*	*OLS*	*ordered logistic*	*logistic*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treatmentTreatment	0.155***	0.095***	0.009	−0.009	0.051	0.022	0.017	0.116
	(0.028)	(0.035)	(0.027)	(0.035)	(0.039)	(0.096)	(0.072)	(0.190)
s1_tsr			0.674***					
			(0.022)					
s1_female					0.118***	−0.099		0.159
					(0.040)	(0.101)		(0.191)
ir_f16_gpa					0.671***	0.551***		1.001***
					(0.029)	(0.064)		(0.106)
Constant	3.462***	3.522***	1.324***	3.483***	0.929***	1.180***		−0.195
	(0.020)	(0.025)	(0.095)	(0.024)	(0.091)	(0.198)		(0.306)
Observations	2,658	2,106	2,106	2,564	2,348	586	2,548	1,941
R^2	0.012	0.003	0.299	0.00002	0.192	0.115		
Adjusted R^2	0.011	0.003	0.299	-0.0004	0.191	0.110		
Log Likelihood								-430.202
Akaike Inf. Crit.								868.405
Note:								*p<0.1; **p<0.05; ***p<0.01

6 Discussion

TODO: discussion Inefficacy of weak measures in the face of divergent student/instructor demographics and massive class sizes. Who perceives similarity What does it mean for relationships and therefore academic success to be built on similarity.

References

Bers, Trudy, and Pam Schuetz. 2014. "Nearbies: A Missing Piece of the College Completion Conundrum." *Community College Review* 42 (3): 167–83. https://doi.org/10.1177/0091552114525834.

Bowen, William G., Matthew M. Chingos, and Michael S. McPherson. 2009. *College: What It Was, Is, and Should Be*. Princeton, U.S.A.: Princeton University Press.

Delbanco, Andrew. 2015. *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton, U.S.A.: Princeton University Press.

Evans, William N., Melissa S. Kearney, Brendan Perry, and James X. Sullivan. 2020. "Increasing Community College Completion Rates Among Low-Income Students: Evidence from a Randomized Controlled Trial Evaluation of a Case-Management Intervention." *Journal of Policy Analysis and Management* 39 (4): 930–65. https://doi.org/10.1002/pam.22256.

Gilbert, Clark G., and Michael B. Horn. 2020. "A Certificate, Then a Degree." *Education Next* 20 (1): 86–87.

Gurantz, Oded, Michael Hurwitz, and Jonathan Smith. 2017. "BOOSTING Hispanic College Completion: Does High-School Recruiting Help More Students Graduate?" *Education Next* 17 (3): 60–67.

Robinson, Carly D., Whitney Scott, and Michael A. Gottfried. 2019. "Taking It to the Next Level: A Field Experiment to Improve Instructor-Student Relationships in College." *AERA Open* 5 (1): 1–15. https://doi.org/10.21105/joss.01686.